Web Appendix:

Lost in Transit: Product Replacement Bias and Pricing to Market

Emi Nakamura and Jón Steinsson

Columbia University

November 25, 2011

A Steps in the Derivation of the Product Replacement Bias Factor

A.1 Decomposing the Regression Coefficients

The OLS estimator of B in equation (6) is

$$B = (\Lambda'\Lambda)^{-1}\Lambda'\Delta p_i,$$

where Λ is a $T \times (\kappa + 1)$ matrix with the Λ_t's as rows and Δp_i is a $T \times 1$ vector with Δp_{it} as its elements.[1] Equation (4) then implies

$$\begin{aligned} B &= (\Lambda'\Lambda)^{-1}\Lambda' \textstyle\int \Delta p_{ik}(s) dkds \\ &= \textstyle\int (\Lambda'\Lambda)^{-1}\Lambda'\Delta p_{ik}(s) dkds \\ &= \textstyle\int B_k(s) dkds \end{aligned}$$

A.2 Deriving Equation (11)

Consider all price spells that end at a particular date. The probability that a given one of these spells is a one period uncensored spell is $f_k(s)(1 - z(s))$. The probability that a given one of these spells is a two period uncensored spell is $f_k(s)(1 - f_k(s))(1 - z(s))^2$, and so on. The overall

[1] For notational convenience, the derivations below assume that all variables have been demeaned and omit having a constant in all regressions. It is well known that a regression on demeaned data yields the same result as the same regression with a constant term on non-demeaned data.

probability that a given one of these spells is uncensored spell is then

$$\sum_{j=0}^{\infty} f_k(s)(1-f_k(s))^j(1-z(s))^{j+1} = \frac{f_k(s)(1-z(s))}{1-(1-f_k(s))(1-z(s))}.$$

Using this we can calculate the average duration of uncensored spells as

$$\frac{1-(1-f_k(s))(1-z(s))}{f_k(s)(1-z(s))}\sum_{j=0}^{\infty}(j+1)f_k(s)(1-f_k(s))^j(1-z(s))^{j+1} = \frac{1}{f_k(s)+z(s)-f_k(s)z(s)}$$

which implies that the frequency of price change in a sample of uncensored spells is

$$f_k(s)+z(s)-f_k(s)z(s).$$

The true frequency of price change of all price spells in the economy is also $f_k(s)+z(s)-f_k(s)z(s)$. Since the both price changes and substitutions have a constant probability of occuring, this implies that the distribution of price spell durations at any point in time is the same for uncencored spells as for all price spells in the economy. This implies that the distribution of price changes is the same at any point in time for uncencored spells as for all price spells in the economy. Thus, the covariance of price changes for uncensored spells with any set of variables — and in particular with Λ — is the same as the covariance of all price changes with that same set of variables.

A.3 Deriving the Fraction of Price Changes that Are the First Price Change for a Product

Since the first observed price change for a product is different from subsequent price changes, we need to know what fraction of price changes are the first observed price change for a product. For a randomly selected price change, the observed "event" preceding this price change is either the product's introduction or another price change. Since these events occur with frequency $\tilde{z}^d(s)$ and $(1-z(s))f_k(s)$, the fraction of measured price changes that are first price changes for a product is $\Phi_k(s) = z(s)/(f_k(s)+z(s)-f_k(s)z(s))$.

A.4 Lifelong Pass-Through as a Special Case

Consider the special case in which firms' optimal prices p^*_{jkt} are a function only of the current exchange rate, and there is no "overreaction" of the first price change, i.e., $\alpha_k(s) = 0$. In this case, $\Delta p^*_{jkt} = B\sum_{\tau=l_j(t)+1}^{t}\Delta e_\tau$, where $l_j(t)$ denotes the time of the previous change in the price of product j before the one at time t and B denotes true pass-through.

The lifelong regression estimates the equation, $\sum_{life} \Delta p^*_{jkt} = B \sum_{life} \Delta e_\tau$, yielding an unbiased estimator of B. Under these assumptions, our bias adjustment equation—equation (10)—simplifies to

$$\begin{aligned} \sum_n B_{nk}(s) &= \tfrac{f_k(s)+z(s)-f_k(s)z(s)}{f_k(s)} \sum_n B^{mm}_{nk}(s) \\ &= (f_k(s) + z(s) - f_k(s)z(s)) \sum_n B^{ch}_{nk}(s) \\ &= (f_k(s) + z(s) - f_k(s)z(s)) B \sum_n n Prob(l_j = n) \\ &= B, \end{aligned}$$

where the first step follows from the properties of an OLS regression, the second step follows from the structural assumption on pricing behavior and the last step follows because the expected duration of price spells in this setting is $1/(f_k(s) + z(s) - f_k(s)z(s))$.

A.5 Deriving the Relationship between True and Measured Frequencies of Price Change and Product Replacement

Here we derive expressions for $\tilde{z}(s)$ and $\tilde{f}_k(s)$ in terms of $f_k(s)$, $z(s)$, and $g(s)$. This allows for the possibility that the observed frequency of price change is lower than the true frequency of price change due to "satisficing" behavior by firms responding to the government's pricing survey as discussed in section 5.2. A product replacement into the economy is measured by the government at time t if the product is accurately observed and a product replacement into the world occurs at time t. A product replacement into the economy is also measured at time t if the product is accurately observed at time t but was not accurately observed a time $t-1$ and a product replacement into the economy occurred at time $t-1$. And so on for earlier periods. This implies that the measured frequency of product replacement is

$$\tilde{z}(s) = g(s)z(s) \sum_{r=0}^{\infty} (1 - z(s))^r (1 - g(s))^r = \frac{g(s)z(s)}{z(s) + g(s) - z(s)g(s)}$$

If the frequency of product replacement in the government's dataset differs from the frequency of product replacement in the world—e.g., because of sample rotation—the expression for $\tilde{z}(s)$ is the same except that $z(s)$ is replaced by the frequency of product replacement in the government's dataset.

To calculate $\tilde{f}_k(s)$, we first calculate the probability that a product is observed in period t and neither a price change nor product replacement has occurred. This is the case if the product was observed at time $t-1$ and no price change or product replacement occurred in period t—an event that has probability $g(s)^2(1 - z(s))(1 - f_k(s))$. It is also the case if the product was last observed in period $t-2$ and has not had a price change or product replacement since—an event that has probability $g(s)^2(1-g(s))(1-z(s))^2(1-f_k(s))^2$; and so on. The total probability of this occurring is thus

$$\frac{g(s)^2(1-z(s))(1-f_k(s))}{1-(1-g(s))(1-z(s))(1-f_k(s))}.$$

Notice that the probability that the product is observed and either a price change or product substitution is observed is then given by

$$g(s) - \frac{g(s)^2(1-z(s))(1-f_k(s))}{1-(1-g(s))(1-z(s))(1-f_k(s))} = g(s)\frac{1-(1-z(s))(1-f_k(s))}{1-(1-g(s))(1-z(s))(1-f_k(s))}.$$

Finally, the probability that the product is observed and a price change is observed is the probability that the product is observed and either a price change or product substitution is observed minus the probability that a product substitution is observed

$$g(s)\frac{1-(1-z(s))(1-f_k(s))}{1-(1-g(s))(1-z(s))(1-f_k(s))} - \tilde{z}(s) = \frac{((g(s)-\tilde{z}(s))f_k(s)}{1-(1-g(s))(1-z(s))(1-f_k(s))}.$$

One must divide this by $1-\tilde{z}(s)$ to get the observed frequency of price change since the denominator in our estimate of the frequency of price change does not include the periods in which a product substitution occurred.

B Multi-Sector Model

For robustness, we have also analyzed product replacement bias in a multi-sector version of our model. In this model, divide product into 15 sectoral groupings of 2 digit HS codes. Within each sector, we assume that the frequency of price change is distributed according to a beta distribution Beta(a, b). To obtain estimates of these parameters, we maximize the log likelihood function presented in appendix C separately for each sector. One interesting fact revealed by this analysis is that there is a large amount of heterogeneity in the frequency of price change *within* the 15 major groups as well as across these groups. We also allow the frequency of product replacement to differ across sectors. Otherwise, the model is identical to the model presented in section 4. We find that this model yields quantitatively similar results to our baseline model. Product replacement bias is slightly larger in the multi-sector model than our baseline model.

C Log-Likelihood in the Presence of Unobserved Heterogeneity in the Frequency of Price Change

We assume that product i has a constant hazard of adjusting, f_i, in each month, where $f_i \sim \text{Beta}(a,b)$. Let us denote the product's lifetime by n_i. These assumptions imply that the total number of price changes in a product's lifetime is distributed according to the binomial distribution, $x_i \sim \text{Bin}(n_i, f_i)$. We assume, furthermore, that f_i is distributed according to the beta distribution, $f_i \sim \text{Beta}(a,b)$.

Given this model, we can write the likelihood of observing a product with length n_i and the total number of price changes x_i as,

$$L = \prod_{i=1}^{I} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} f_i^{a-1}(1-f_i)^{b-1} \binom{n_i}{x_i} f_i^{x_i}(1-f_i)^{n_i-x_i} \quad (1)$$

$$= \prod_{i=1}^{I} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} f_i^{x_i+a-1}(1-f_i)^{n_i-x_i+b-1} \binom{n_i}{x_i} \quad (2)$$

We can integrate out the $f_i's$ to get,

$$L = \prod_{i=1}^{I} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \binom{n_i}{x_i} \frac{\Gamma(a+x_i)\Gamma(b+n_i-x_i)}{\Gamma(a+b+n_i)}. \quad (3)$$

The log-likelihood function is, therefore,

$$\log L = n\log\Gamma(a+b) - n\log\Gamma(a) - n\log\Gamma(b)^n + \sum_{i=1}[\log n_i! - \log x_i! \quad (4)$$

$$- \log(n_i - x_i)! + \log\Gamma(a+x_i) + \log\Gamma(b+n_i-x_i) - \log\Gamma(a+b+n_i)]. \quad (5)$$

D Derivation of Equations (16) and (17)

Under the simplifying assumptions that the frequency of price change and product substitution are constant over time for each product group and that long-run pass-through is the same for all products, equation (15) simplifies to

$$\sum_n B_n^{mm} = \int \frac{f_k}{f_k + z - f_k z}\left[\Phi_k(1+\alpha_k)\sum_n B_{nk} + (1-\Phi_k)\sum_n B_{nk}\right] dF(k).$$

Manipulation of this equation and equation (13) yields

$$\sum_n B_n^{mm} = \int \frac{f_k}{f_k + z - f_k z} dF(k) \sum_n B_n$$
$$+ z\int \frac{f_k}{f_k + z - f_k z}\left[(1+\alpha_k)\sum_n B_{nk}^{ch} - \sum_n B_{nk}^{ch}\right] dF(k).$$

Since $f_k/(f_k + z - f_k z) < 1$, this equation implies that

$$\sum_n B_n^{mm} < \int \frac{f_k}{f_k + z - f_k z} dF(k) \sum_n B_n + z \left[\int (1 + \alpha_k) \sum_n B_{nk}^{ch} dF(k) - \int \sum_n B_{nk}^{ch} dF(k) \right].$$

A conservative measure of product replacement bias is thus given by

$$\sum_n B_n^{mm} = \int \frac{f_k}{f_k + z - f_k z} dF(k) \sum_n B_n + z \left[\int (1 + \alpha_k) \sum_n B_{nk}^{ch} dF(k) - \int \sum_n B_{nk}^{ch} dF(k) \right].$$

The last term in this expression reflects the correction for "overreaction" of the first price change discussed in the body of the paper. Empirically, we estimate the size of this term by comparing the sum of the coefficients for a regression of the first observed price change for all products with two or more price changes on lagged exchange rate changes with the sum of the coefficients for a regression of the second observed price change for all products with two or more price changes on lagged exchange rate changes. Since this comparison is based on the sample of products with two or more price changes (recall that B_{nk}^{ch} is the regression coefficient for all price changes after the first price change), we need to adjust for the fact that the frequency of price change is higher for this subsample of products than the population as a whole. Specifically, $\int (1 + \alpha_k) \sum_n B_{nk}^{ch} dF(k)$ is smaller in this subsample by a factor equal to the ratio of the average duration in the subsample of products with two or more price changes relative to the average duration in the population as a whole. Applying this adjustment and assuming that α is the same accross products with different frequencies of price change, yields

$$\sum_n B_n^{mm} = \int \frac{f_k}{f_k + z - f_k z} dF(k) \sum_n B_n + z \frac{\bar{d}}{\bar{d}_2} \left[(1 + \alpha_k) \sum_n B_{2n}^{ch} - \sum_n B_{2n}^{ch} \right], \qquad (6)$$

where $\bar{d}$ denotes the average duration of all price spells, $\bar{d}_2$ denotes the average duration of all price spells of products with two or more price changes, and the subscript "2" in B_{2n} indicates that we are calculating the sum of the coefficients for products that have two or more price changes. If pass-through is higher for products with a higher frequency of price change this adjustment will be smaller, since the ratio of $\int (1 + \alpha_k) \sum_n B_{nk}^{ch} dF(k)$ will be less than the ratio of average lengths. Equation (6) therefore presents a very conservative (i.e., lower bound) estimate of the magnitude of product replacement bias.

E Product Replacement Bias and BLS Price Imputation

As we discuss in section 2, transaction prices are missing in about 40% of the product-months in the IPP dataset. During these periods, IPP uses various imputation procedures to "connect the dots" between reported prices. The primary method used by the BLS to impute prices between periods when it gets a new price quote is to "carry forward" the last observed price (the method we use in our empirical analysis). In some cases, however, the BLS uses other imputation procedures including linear interpolation and cell mean imputation methods (see Feenstra and Diewert, 2000 for more details). Below, we discuss the robustness of the method we use in our empirical analysis to alternative imputation procedures.

All the BLS imputation procedures merely "fill in the dots" between the observed prices for short periods when prices are not observed. Any price change that is introduced as a part of the imputation procedure is reversed as soon as a new transaction price is observed. As a consequence, any adjustment to exchange rates that such imputed price changes may contain are reversed and do not affect long-run pass-through.

The "fill in the dots" nature of BLS imputation implies that whatever such imputation method the BLS uses, one can produce an alternative index using a "carry forward" imputation method and this alternative index will yield the same measured long run pass-through as the actual BLS series. To verify this we have constructed an alternative index based solely on the "carry forward" imputation method and find that it yields identical estimates of long-run pass-through to the official aggregate BLS index, and in addition tracks the official index closely (except at very high frequencies).

Since both the actual index and the alternative index based solely on "carry forward" imputation yield the same measured log-run pass-through, one can formulate an adjustment for product replacement bias based on either one. The theoretical adjustment we present in section 4 is formulated for the alternative "carry forward" index in that prices are assumed to remain unchanged whenever firms fail to report prices accurately (see section 4.3). This implies that the frequency of price change concept that appears in our adjustment for product replacement bias is the frequency of price change for the carry forward index. This is why, in section 5, we use a "carry forward" procedure when calculating the frequency of price change that we input into our estimate of the product replacement bias factor.

Intuitively, product replacement bias arises because some movements in exchange rates are "unaccounted for" by subsequent price changes when the price series is prematurely truncated by a substitution. Imputation procedures such as cell-mean imputation may lead to additional price changes between existing observed prices—and slight differences in high frequency dynamics of the resulting price index—but they will never add any additional long-run "responsiveness" to exchange rates to the series or affect long-run pass-through, since price changes associated with imputations are always subsequently reversed when the products price is again observed. Thus, any exchange rate movements that are "unaccounted for" in a "carry-forward" index will still be unaccounted for in an index based on cell-mean imputation or other methods of imputation. The crucial statistic for adjusting for product replacement bias is the fraction of time that is unaccounted for because it belongs to the last observed price change of a product. One way to calculate this is based on the frequency of price change estimated from a carry forward index.

References

Feenstra, R. C., and E. W. Diewert (2000): "Imputation and Price Indexes: Theory and Evidence from the International Price Program," Working Paper, UC Davis.